Journal of Information Technology Management
Print ISSN: 2008-5893
Online ISSN: 2423-5059

Hybrid Bio-Inspired Clustering Algorithm for Energy Efficient Wireless Sensor Networks

Amirhossein Barzin

PhD Candidate, Industrial Engineering, Azadi Pardis of Yazd University, Yazd University, Yazd, Iran. E-mail: barzin@stu.yazd.ac.ir

Ahmad Sadegheih

Corresponding author, Professor, Industrial Engineering, Faculty of Industrial Engineering, Yazd University, Daneshgah Blvd., Safayieh, PO Box: 89195-741, Yazd, Iran. sadegheih@yazd.ac.ir

Hassan Khademi Zare

Professor, Industrial Engineering, Faculty of Industrial Engineering, Yazd University, Daneshgah Blvd., Safayieh, PO Box: 89195-741, Yazd, Iran. E-mail: hkhademiz@yazd.ac.ir

Mahboobeh Honarvar

Assistant Professor of Industrial Engineering, Faculty of Industrial Engineering, Yazd University, Daneshgah Blvd., Safayieh, PO Box: 89195-741, Yazd, Iran. E-mail: mhonarvar@yazd.ac.ir

Abstract

In order to achieve the sensing, communication and processing tasks of Wireless Sensor Networks, an energy-efficient routing protocol is required to manage the dissipated energy of the network and to minimalize the traffic and the overhead during the data transmission stages. Clustering is the most common technique to balance energy consumption amongst all sensor nodes throughout the network. In this paper, a multi-objective bio-inspired algorithm based on the Firefly and the Shuffled frog-leaping algorithms is presented as a clustering-based routing protocol for Wireless Sensor Networks. The multi-objective fitness function of the proposed algorithm has been performed on different criteria such as residual energy of nodes, inter-cluster distances, cluster head distances to the sink and overlaps of clusters, to select the proper cluster heads at each round. The parameters of the proposed approach in the clustering phase can be adaptively tuned to achieve the best performance based on the network requirements. Simulation outcomes have displayed average lifetime improvements of up to 33.95%, 32.62%, 12.1%, 13.85% compared with LEACH, ERA, SIF and FSFLA respectively, in different network scenarios.

Keywords: Wireless Sensor Networks; Clustering; Bio-inspired Algorithm; Firefly Algorithm; Shuffled Frog Leaping Algorithm.

DOI: 10.22059/jitm.2019.280639.2354

1. INTRODUCTION

Rapid enhancements in micro-electro-mechanical systems (MEMS) along with wireless communication technology have caused wireless sensor networks (WSN) to become one of the key technologies in the current century due to their large amount of applications. WSNs consist of a large number of low-cost, low-power tiny sensing nodes, randomly dispersed in a target area far from human reach without specific infrastructure. Depending on the sensor installed in the node, a particular occurrence i.e. pressure, temperature, humidity is sensed by the sensor unit and transmitted to electrical signals then is sent to the base station (called sink) via the radio interface. The sink is located between the user and the network and is responsible for gathering information from the nodes. WSNs have various applications in different areas and researchers have developed many techniques to improve their performance in an application-specific way. The WSNs' aspects such as the speed of operation, computation, fault tolerance, network autonomously and control, are applied to identify and track adjacent hostile targets in military applications (Sohraby, 2007).

Other applications of WSN's in healthcare systems, industry and agriculture surveillance, earthquake detection, product quality monitoring, and remote area control, have been effectively tried out (Sohraby, 2007). Ko, Lau, and Sham (2008) presented the scheme and application of a distributed wireless sensor network to track a kind of rescue robot that searches for the heat of the active human body using its thermal array sensor. By analyzing the signal strength, the wireless sensor network aids to path the location of the robot. Pantoni and Brandão (2013) proposed a simple, reliable gradient-based routing protocol for WSNs, in which, the authors' concentration was to implement an efficient routing protocol based on the requirements for a street lighting application. Minaie and Sanati-Mehrizy (2013) have addressed some of the medical care system applications of WSNs. Where authors divide these applications into three main categories including patients' monitoring in the clinical sceneries, surveillance of the Home and elderly care centers for chronic and old patients, and long-standing databases' collections for clinical data of health applications. Sensors can either monitor the patient in the network or help the person with a disability. Deployment of WSNs faces many challenges, such as energy restrictions, lifetime longevity, security, communication reliability, design, etc. (Fanian & Kuchaki, 2019).

It should be noted that, due to the conflicting essence among these challenges, it is hard to balance all these aspects. Recently, several techniques have been offered concerning one of the main challenges in WSNs; the energy restrictions of sensor nodes, to minimize the energy dissipated by the nodes leading to network lifetime longevity. These methods include data gathering, data correlation, energy harvesting, optimal deployment, beamforming, resource allocation using cross-layer design, sleep-wake scheduling, mobile relays and sinks, clustering and multi-hop routing (Zahedi et al., 2018). In this paper, the authors focus on the clustering and routing in WSNs.

1.1 MOTIVATION

WSNs' technology is employed as the basic infrastructure of the Next Generation Networks (NGN) including Internet of things (IoT), Sensor Control Networks (SCNs), Ubiquitous Sensor Networks (USNs), Machine-Oriented Communications (MOC) and so on. While deploying WSNs, the network designer is required to consider many issues, which involve in selecting between several alternatives. These issues include deciding the network topology, the number of sensor nodes, cluster head selection, security model, the relative position of elements, hardware and software for both sensor nodes and servers. The ultimate goal of these choices is to make the WSN solve problems set for it, effectively (Butenko et al. 2014).

Bio-inspired algorithms, on the other hand, are soft computing techniques, which have been widely pondered to solve a broad range of optimization problems. For example, a genetic algorithm (GA) was employed to improve the efficiency of construction automation system (Wi et al., 2012). Likewise, particle swarm optimization (PSO) was applied to solve many optimization problems in manufacturing (Issam et al., 2013; Thitipong & Nitin, 2011). Similarly, clustering and routing are two renowned optimization problems, which are researched broadly for developing many bio-inspired based algorithms in the field of wireless sensor networks (Pratyay & Prasanta, 2014). Structurally, in WSNs the sensor nodes are small and have often limited irreplaceable energy supply. They send information at short distances. Thus, innovative techniques that minimize energy consumption and maximize the life span of the network are significant (Pratyay & Prasanta, 2014). Clustering-based routing is one of the prevailing energy-efficient routing techniques. In this scheme, to collect data, sensor nodes are segmented into non-overlapping clusters. Each cluster has a cluster head (CH). The sensor nodes belonging to each cluster send their data only to their CH. The CH compresses and integrates (aggregates) correlated data and sends it to the sink (Heinzelman, Chandrakasan, & Balakrishnan, 2000). Clustering-based routing protocols include the setup and steady phases. During the setup phase, CHs are selected and each node is connected to the nearest CH. A steady (communication) phase, to reduce the need for additional and unnecessary data transfers, CHs do data aggregation (Heinzelman et al., 2000).

Clustering reduces the length of the routing table, which is stored at the specific nodes via restricting the route formed up inside the cluster. Moreover, clustering can preserve the bandwidth by limiting the domain of inter-cluster communications to CHs and circumvents unnecessary messages exchanges amongst sensor nodes. In homogeneous networks, if the collected data packets are sufficiently correlated, they will be aggregated and will be sent to the sink via a single- or multi-hop scheme. Whereas in heterogeneous wireless sensor networks, normally, a large number of low-cost nodes do the sensing, while a few ones having relatively more energy perform data fusion, filtering, and transportation. In large-scale networks, clustering supports scalability, multi-gateway topologies as well as data aggregation at cluster heads. The Cluster Head can prolong the life expectancy of battery for each sensor leading to network lifetime durability through implementing optimized management strategies. For small-scale networks, the participation of fewer nodes in the data transmission can provide efficient use of energy resources. However, devolving more tasks to CHs will require more energy consumption to process and transmit each cluster's data, which will result in premature and irregular network depletion (Pratyay & Prasanta, 2014).

Generally, clustering algorithms presented for WSNs in the literature can be categorized based on the network architecture, operation model and the objective of the node grouping process including the desired count and properties of the generated clusters. Some of the attributes used to classify clustering algorithms are cluster properties, cluster-head capabilities, and clustering process. Each of these attributes has its sub-categories. Abbasi and Younis, (2007) presented a comprehensive taxonomy of the clustering algorithm in WSNs. Besides, all the attributes affecting the design of clustering algorithms, there are several supplementary challenges, which influence the design of routing protocols. Some of them can be outlined as follows (Fanian & Kuchaki, 2019):

- *Energy constraints*: Each sensor has a battery-powered with limited energy, which turns the computing, sending and receiving of data into a challenging task, so the lifetime of sensors highly depends on it.
- *Data aggregation*: Adjacent sensor nodes in the homogeneous networks may sense duplicate events. In this case, a proper aggregation method can be utilized to prevent redundant data from being transmitted to the sink.
- *CH selection*: The mechanism of selecting CHs is of the highest significance in both clustering-based and multi-hop routing protocols, as only the selected CHs contribute in the multi-hop routing.
- *Time complexity*: Time is a major issue in the convergence of clustering algorithms. As can be reviewed in Table-1 Some of the proposed clustering algorithms such as LCA, RCC and

CLUBS, have convergence time of O(n), in which n represents the number of sensor in the network. Therefore, these types of clustering algorithms are suitable to the networks with small number of nodes. Hitherto, convergence time has improved dramatically in some more current algorithms like in LEACH, HEED and MOCA. Hence, they are apt for networks having large number of nodes. Generally, variable convergence time algorithms enable more control of the cluster properties than the constant time ones.

Table 1 compares some existing clustering-based algorithms based on the most important clustering attributes.

Table 1. Comparison of some clustering algorithms for WSNs

Clustering approaches	Time Complexity	Node mobility	Cluster overlapping	Location awareness	Energy efficient	Failure recovery	Balanced clustering
LCA	Var. O(n)	Possible	No	Needed	No	Yes	OK
Adaptive clustering	Var. O(n)	Yes	No	Needed	N/A	Yes	OK
CLUBS	Var. O(n)	Possible	High	Not Needed	N/A	Yes	OK
RCC	Var. O(n)	Yes	No	Needed	N/A	Yes	Good
LEACH	Const. O(1)	Fixed BS	No	Not Needed	No	Yes	OK
HEED	Const. O(1)	Stationary	No	Not Needed	Yes	N/A	Good

Therefore, while efforts to decrease the dissipated energy have enclosed the different aspects of WSNs, many important objectives persist untouched, some of these concerns include:

- There is no all-purpose approach to determine and optimize the dissipated energy.
- Current approaches focus on one feature and may load energy consumption in other aspects.
- Existing approaches miss quantitative measures of energy consumption of the entire network.
- NP-hardness of the energy efficiency and the lifetime longevity problem in WSNs approves metaheuristic approaches.

1.2 OUR CONTRIBUTION

In this paper, the authors suggest a new adaptive clustering algorithm for WSNs called FSFA. This algorithm combines two population-based meta-heuristics, the shuffled frog-leaping algorithm (SFLA) and the Firefly algorithm (FFA), in a high-level hybridization way (Al-Ghazzali 2009). Roulette wheel selection (RWS) and some local improvements in the setup phase help our hybrid approach to converge faster comparing with each of the basic algorithms (Al-Ghazzali. 2009). Key contributions to this paper are listed as follows:

- Developing a hybrid bio-inspired based metaheuristic algorithm (called FSFA) and employ it as a clustering protocol in WSNs.
- Introducing a new adaptive application-specific multi-objective fitness function, which can be adjusted based on the application scope.
- Considering multiple criteria (e.g., inter- and intra-cluster distances, residual energy of

nodes, distances from the sink, estimated energy consumption, overlap and load of clusters) to select appropriate the cluster heads. The relative significance of these criteria can also be tuned according to the application specifications.

- Performing FSFA in different scenarios to demonstrate its performance against the existing protocols, in terms of energy consumption and network lifetime.

2. RELATED WORKS

Recently, several clustering-based routing protocols have been presented to deal with energy efficiency and the lifetime longevity in WSNs. These techniques can be distinguished by the mechanism of CH-selection at each round (Jalali et al., 2015). In some approaches, CH selection is considered as a decision-making problem under particular criteria (Butenko et al., 2014). Optimal CH selection in some other approaches deliberated as to be NP-hard. Generally, solution-searching techniques for NP-hardness can be classified into three categories including exact (complete), heuristic and random search techniques (Al-Ghazzali 2009). In the exact technique, at first, all possible solutions are generated and evaluated. Then, the best one is selected. Hence, the optimal solution is usually attainable. However, since clustering and routing in WSNs inherit the complication of the NP-hard problems, exact solutions cannot achieve through a polynomial time complexity, even for small-size networks. Accordingly, researchers are addressing this problem through heuristic and random search methods. Commonly, classical and fuzzy-based approaches belong to the heuristic search category, while bio-inspired algorithms are known as random search methods. From the time complexity point of view, classical and fuzzy-based techniques are performed in less time; however, metaheuristic-based methods can achieve better performance, as they effectively investigate the entire search space (Al-Ghazzali 2009). In this section, the existing clustering-based routing protocols are grouped in classical, fuzzy-based, metaheuristic-based and hybrid approaches.

2-1 CLASSICAL APPROACHES

The Low Energy Adaptive Clustering Hierarchy (LEACH) is one of the most renowned clustering-based protocols used in WSNs introduced by Heinzelman, Chandrakasan, and Balakrishnan (2002). LEACH is a distributed protocol, wherein nodes freely decide to become a CH. In this protocol, each node might be qualified to be a CH only once in $1/p$ consecutive rounds, where p is a preferred percentage of CHs. In order to prevent energy drainage in CHs, LEACH vigorously switches the workload of CHs amongst the nodes. The operation of LEACH at each round includes the setup phase and the steady-state phase. In the first phase, every node n selects a random-generated number in the range of [0,1]. It is eligible to be a CH in the current round if the corresponding random number is smaller than the threshold value denoted by $T(n)$ as formulated in Eq. (1). In which p denotes the preferred percentage of CHs, r is the current round, and G represents the set of nodes that have not functioned as a CH in the previous 1/p rounds.

$$(1) \qquad T(n)=\begin{cases} \dfrac{p}{1-p\times(r \bmod 1/p)} & if \;\; n\in G \\ 0 & \text{otherwise} \end{cases}$$

When a node has been qualified as a CH, it advertises a message to the rest of the nodes. Amongst all non-CH nodes who are listening, the one who receives the strongest announcement signal becomes a member of the CH. Once all clusters are shaped, the next phase begins. In the steady-state phase, the network functions are done within several time slots. In each time slot, CH generates a time division multiple access (TDMA) scheme to assign a time slot for each member to send its data. Once all member nodes transmit their data to the analogous CH, the CH aggregates the combination of data into a single packet and sends it towards the sink. Afterward, the current round ends and a new round begins. The key weakness of LEACH is that some particulars such as position and the remaining energy in the nodes have not been accounted for CH selection in clustering. Moreover, LEACH assumes that the CHs send the gathered packets to the sink directly by single-hop communication. This assumption causes the network to run out of energy sooner and is unreliable for large scale WSNs. Kassan, Gaber, and Lorenz (2018) introduced a novel method via combining a non-cooperative Game Theory (GT) method using a decentralized clustering algorithm to deal with the problem of the network lifetime longevity. This approach employed the GT techniques to reduce the energy dissipated by the WSN, via decreasing the number of forwarded packets leading to the network lifetime improvement. The simulations outcomes showed that the proposed approach outperformed existing distributions based clustering algorithms without GT, such as LELC and LEACH in terms of saving energy and increasing the number of data packets received by the sink. Energy-aware Routing Algorithm (ERA) presented by Amgoth and Jana (2015) comprises two phases including clustering and routing phases. During the first phase, prior to nodes' competition to become a CH, each node is assigned an autonomous timer. This timer represents the maximum time determined for selecting CHs. The higher the energy level, the shorter the time slot to let the node be selected as a CH. Once the timer was over and no messages were received from other CHs, the node would advertise itself as a CH via disseminating an advertisement within a specific range. Otherwise, if a node receives a message from a CH before the timer's run out, it turns into a non-CH node.

An AHP clustering approach was offered by Hanifi, Taghva, Haghi and Feizi (2018) in which, in the first phase of the clustering, the position of the nodes were sought using two identified positions including the sink's position and two assumed nodes out8of the area of interest. In the second phase, the authors determine the CHs based on the criteria including the remaining energy, the distance of the nodes from the cluster head, the distance of CHs from the base station, the number of neighbors and the centrality, using the multi-criteria decision-making method. The proposed method was simulated in the NS2 simulator and its outcomes were evaluated and compared with the existing methods including NEECP E-LEACH protocols. The authors showed that the proposed method improved the energy consumption, the network life span, the average packet delivery, and the average delay. A two-level TOPSIS based clustering scheme for WSNs was proposed by Hamzeloei and Khalily (2016) in which the cluster head selection was performed by a multiple criteria decision-making method regarding four criteria. These criteria include the residual energy of nodes, the number of neighbors, the distance below the sink and the transmission range for each sensor node. The proposed model‘s results were compared with the AHP and LEACH clustering approaches and showed improvement in terms of network lifetime.

2-2 FUZZY APPROACHES

Owing to uncertainties occurring in the WSN environment and overlapping parameters affecting the role of CHs, some protocols make use of the fuzzy logic. Using fuzzy variables, the inherent uncertainties of WSNs can be handled. Moreover, using fuzzy logic instead of the classical formula in selecting CHs is more flexible. Some of these techniques are discussed in this section. Ran, Zhang, and Gong, (2010) introduced LEACH with Fuzzy Logic (LEACH-FL). This technique is a fuzzy version of LEACH, which qualifies appropriate CHs, by applying the Mamdani fuzzy inference system. To design the fuzzy inference system of this protocol, three fuzzy variables are chosen as inputs: energy level, density, and distance of nodes from the sink. At each round, first nodes are sorted into a descending order according to the fuzzy output. Then, those nodes, which have the highest output, are qualified as CHs. All calculations associated with the fuzzy qualification of CHs are executed over a central processor located in the sink.

Bagci and Yazici (2013) suggested another fuzzy-based clustering algorithm named EAUCF. The main objective of this approach is to reduce the workload of those clusters, which are located close to the sink or have low energy levels. In EAUCF, a random number in the range of 0 and 1 is assigned to each node. A node becomes a CH, only if the assigned number is smaller than a pre-determined threshold. A fuzzy inference system founded on the distance to the sink and the residual energy of the nodes is employed to calculate the competitive radius of the conventional CHs. Once the competitive radius for conventional CHs is determined, each CH contests with other CHs in the radius. As long as a conventional CH does not receive a message indicating a higher level of energy in the radius, it remains a CH.

2-3 METAHEURISTIC APPROACHES

The literature review illuminates that the lifetime longevity of the wireless sensor networks is an NP-hard problem (Pratyay & Prasanta, 2014). Accordingly, researchers are addressing these problems by utilizing meta-heuristics. LEACH-C proposed by Tripathi, Gaur, Laxmi, and Battula (2013) is a centralized metaheuristic-based energy-aware extension of LEACH. In this scheme, for the setup phase of the first round, each node sends information containing its residual energy, location along with the sensed data to the sink. Then, at the setup phase of the next round, the sink qualifies some nodes as the current CHs via simulated annealing (SA) algorithm. The CH candidates whose energy levels are higher than the average residual energy of all active nodes, are considered to be finally qualified as CHs via simulated annealing. After the CH-selection procedure, the sink advertises IDs of selected CHs in the network to inform other nodes of CHs selected for the current round. Other operations in LEACH-C are similar to those in the LEACH. Abba Ari, Yenke, Labraoui, Damakoa, and Gueroui (2016) presented a novel cluster-based routing protocol called ABC-SD. The proposed scheme is based on the nature-inspired efficient and rapid searching aspects of the Artiifcial Bee Colony (ABC) metaheuristic to build up low power consumption clusters. A multi-objective fitness function has been formulated, to choose the proper CHs. The centralized location-unaware clustering algorithm is executed at the sink, using energy levels and the adjacency information of the sensor nodes as input parameters. The proposed protocol has been massively applied to different topologies in various network scenarios. Outcomes were compared with other renowned bio-inspired based cluster-based routing protocols. The obtained results outperformed the rivals in terms of network lifetime, network coverage and the number of packets delivered at the sink. Jabeur (2016) proposed a firefly-based clustering approach including two clustering phases. During the micro-clustering phase, sensors were grouped in clusters autonomously. During the second phase (macro-clustering), clusters refined through a competition to aggregate small nearby clusters. Simulations show promising results where the number of clusters tends to stabilize independently from the density of the network and the various communication ranges of sensors. Anandamurugan and Abirami (2017) suggested another meta-heuristics based approach for clustering in WSNs with the anti-predator ability, which evades the algorithm from being trapped in the local optima. Results approved local optima evading if compared with original SFLA and particle swarm optimization in clustering of WSNs.

2-4 HYBRID APPROACHES

In this sub-section few, lately proposed, hybrid clustering schemes will be briefly discussed. Zahedi, Akbari, Shokouhifar, Safaei, and Jalali (2016) suggested swarm Intelligence Fuzzy (SIF). SIF is a centralized clustering protocol that utilizes the Mamdani fuzzy system for the selection of CHs at each round. This scheme uses three input variables: the energy level, the distance to the cluster centroid, and the distance to the sink. At first, all nodes are clustered via the fuzzy c-means algorithm, then, in each cluster, one node is selected as CH by the Mamdani fuzzy inference system. Gupta and Jha (2018) offered a novel energy-balanced clustering protocol based on improved cuckoo search. It uses an objective function for uniform distribution of CHs throughout the network. Additionally, an improved harmony search based routing protocol is proposed for the multi-hop routing of the data packets from CHs towards the sink. Fuzzy SFLA (FSFLA) was offered by Fanian and Rafsanjani (2018) which employs the SFLA algorithm together with the fuzzy inference system in a central CPU at the sink to select the CHs for single-hop schemes. It considers the number of neighboring nodes, the remaining energy, the distance from the sink and the history of nodes in the CH-selection procedure (Fanian and Rafsanjani, 2018).

3. OVERVIEW ON BASIC ALGORITHMS

In the following sections, the basic bio-inspired based clustering algorithms used in the proposed model are briefly discussed.

3.1 FIREFLY ALGORITHM

The first clustering approach is based on the firefly algorithm (FFA). FFA was inspired by the natural behavior of fireflies' swarm and was first introduced by Xin-She Yang et al. (2008). Fireflies work based on the phenomenon of bioluminescence; the process in which Firefly insect yields flashes of short duration (Mukhdeep & Singh, 2016). In this approach, the intensity of flash is a vigorous parameter. There are three rules to follow with the firefly approach: Firstly, the attractiveness of each firefly is independent of its gender. Secondly, the tendency for fireflies to be attracted to each other depends on the brightness of the flash. Thirdly, the brightness of the fireflies is determined using the objective function. The brightness of flash is termed as the attractive factor. This factor depends on the intensity of light. The most commonly used variables of this algorithm are attractiveness factor and light intensity (Fister et al., 2013). In FFA, at first, a population of fireflies is randomly generated. In the next step, the fitness evaluation and population updating are performed iteratively, until the maximum number of iterations is attained. In the population-updating phase, the movement of firefly *i* towards another more attractive (brighter) firefly *j* is done according to Eq. (2):

$$x_i = x_i + \beta e^{-\gamma d_{ij}^2}\left(x_j - x_i\right) + \alpha\left(r - 0.5\right) \qquad (2)$$

Where the first term is the present location of the firefly *i*, the second term pertains to the attractiveness of the firefly *j*, and the third term is a random displacement. d_{ij} is the Euclidean distance between fireflies *i* and *j*, *r* is a random number in the range of [0,1], and α, β, and γ are three constant parameters (Fister et al., 2013).

3.2. SHUFFLED FROG-LEAPING ALGORITHM

The second clustering approach is founded on another bio-inspired based meta-heuristic, which was initially introduced by Eusuff, and Lansey (2006) named shuffled frog-leaping algorithm (SFLA). The SFLA principally developed as a population-based metaheuristic capable of performing a cognizant heuristic search through any mathematical function seeking solutions for combinatorial optimization problems. It can be employed to solve several complicated optimization problems, with the essence of the non-linearity, non-differentiability and multi-modality (Khorsandi et al., 2011). It associates the advantages of the genetic-based memetic algorithm with the social behavior-based PSO algorithm (Khorsandi et al., 2011). In this algorithm, a virtual population of possible solutions is generated by a set of frogs (solutions), which segmented to dissimilar subgroups called memeplexes, by each a local search is performed. Inside each memeplex, the individual frogs hold ideas that might be affected by the ideas of other frogs. After a predefined number of memetic evolution paces, ideas are passed among memeplexes through a shuffling process. The local search and the shuffling process continue until the defined convergence criteria are satisfied (Khorsandi et al., 2011). The SFLA can be briefly described thru the following steps:

1. Initialization: Select *m* and *n*, where *m* represents the number of memeplexes and *n* denotes the number of frogs in each memeplex. Thus, the total sample size, F would be $F = m \times n$.
2. Population generation: An initial population of F frogs is created random uniformly. For a *S*-dimensional problem, each frog *i* is represented by *S* variables, such as $F_i = (f_{i1}, f_{i2}, \ldots, f_{iS})$.
3. Sorting and distribution: Frogs are first sorted in descending order based on their fitness values, then the entire population is grouped in *m* memeplex, each contains *n* frogs. In this process the first frog moves to the first memeplex, the second frog moves to the second , frog *m* moves to the m_{th} memeplex, and frog $m+1$ moves back to the first memeplex, and so on.
4. Memeplex evolution: This step is based on a local search, where in each memeplex, frogs with the best and the worst fitness are identified as F_b and F_w respectively. The frog with the global best fitness is identified as F_g separately. Then, an evolution process is applied using Eq. (3) and (4) to improve the position of the frog having the worst fitness.

$$D_i = rand() \times \left(F_b - F_w\right) \quad (3)$$

$$F_w^{i+1} = F_w^i + D_i \quad , \left(D_{\min} \le D_i \le D_{\max}\right) \quad (4)$$

   Where rand () is random number between zero and 1, D_{min} and D_{max} are the minimum and the maximum displacement allowed in a frog's position respectively. If this process generates a better solution, it replaces the worst frog; else, if Eq. (3) and (4) do not improve the worst solution, F_b of Eq. (3) is replaced with F_g adjusting to Eq. (5):

$$D_i = rand() \times \left(F_g - F_w\right) \quad (5)$$

5. Shuffling: after a predefined number of memeplex evolution steps, all frogs of memeplexes are collected, the new population is sorted in descending order according to fitness.

6. Stop condition: if a global solution or a fix number of iteration reached, the algorithm stops. Otherwise, jumps to step (2) and repeat again.

4. SYSTEM MODEL

In this section, the system model including energy, network and lifetime models are discussed respectively.

4.1 ENERGY MODEL

This paper has employed an energy model named “the first order radio model” (Oladimeji et al., 2017). In this scheme, the transmitter dissipates the energy of *ETx(l,d)* to manage the radio electronics and the power amplifier, whereas the receiver dissipates the energy of *ERx(l)* when managing the radio electronics, as shown in Figure 1. Depending on the distance (*d*) between the transmitter and receiver, the free space (*d*2 power loss) and the multipath channel fading (*d*4 power loss) models were used for all the try-outs carried out.

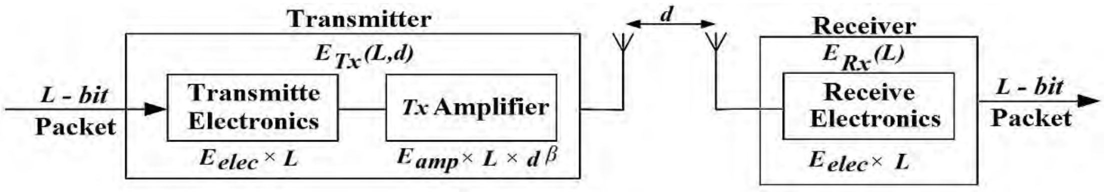

Figure 1. Radio Energy Dissipation Model.

The power-amplifier is properly managed so that if the distance were less than a threshold distance d_0, the free space (*fs*) model would employ elsewhere, and the multipath (*mp*) model would be used. Therefore, for transmitting an *L-bit* message to a distance d, the radio spends:

$$E_{Tx}(L,d) = E_{Tx_{elec}} + E_{Tx_{amp}} = \begin{cases} L \times E_{elec} + \varepsilon mp \times L \times d^4, \ if \ \ \mathrm{d} \succ \mathrm{d}_0 \\ L \times E_{elec} + \varepsilon fs \times L \times d^2, \ if \ \ \mathrm{d} \prec d_0 \end{cases} \quad (6)$$

Furthermore, to receive *L-bit* message, the radio uses energy calculated as:

$$E_{Rx}(L,d) = L \times E_{elec} \quad (7)$$

Where $d_0 = \sqrt{\varepsilon fs / \varepsilon mp}$ denotes the threshold distance. The electronics energy denoted by E_{elec},is effected by factors such as the digital encoding, modulation used as well as filtering type, and spreading of the signal effect. On the contrary, the amplifier energy, *εmp* or *εfs* depends on the distance to the receiver and the acceptable bit-error rate. *β* determines the dissipated energy according to the distance between transmitter and receiver (Oladimeji et al., 2017).

4.2 NETWORK MODEL

According to the application research background, the following network assumptions are considered in the modeling of the proposed algorithm (Oladimeji et al., 2017; Singh et al., 2017):

- All sensor nodes are homogeneous and randomly distributed with few CHs. Once they are deployed, they become stationary, with constant initial energy. Battery recharge and replacement are almost impossible for the entire operation.
- Location of the sink is fixed. It could be placed either inside, or outside of the sensing field, depending on the network scenario.
- To determine the location, all sensor nodes have been embedded with positioning devices (i.e. GPS).

- The communication channel is assumed to be symmetric. That means the energy required for transmitting data from sensor node s_1 to sensor node s_2 is equal to the energy needed to transmit a message from node s_2 to node s_1 for a particular signal to noise ratio (SNR).
- This scheme supports TDMA protocol providing for MAC layer communication. CHs use slotted carrier-sense multiple access (CSMA) MAC protocol to communicate with the sink.
- Date aggregation is done assuming that each CH collects the data from its member nodes and aggregates it into a single packet of fixed length regardless of the number of received packets. Nodes, close to each other, have correlated data.

4.3 LIFETIME MODEL

Perhaps the most important metric for performance evaluation of wireless sensor networks is the network lifetime. Regardless of how the network lifetime is defined, it strongly depends on the lifetimes of each single nodes building up the network. The lifetime modeling of a single sensor node depends on two factors: how much energy it consumes over time, and how much energy is available for its later use. Various definitions of the network lifetime are given in the literature (Dietrich & Dressler, 2009).

It is important to understand that the network lifetime definition is rather reliant on the specific application, and there is no exact definition suitable and applicable for all applications (Zahedi et al., 2016). For instance, in a medical surveillance network, lifetime is determined by the time the first mode is frozen and is not able to keep on data transition or shortly the First Node Dies (FND). Accordingly, in such a case the information gathered from all sensor nodes need to be distinguished and perishing the data of a single sensor node may generate irreparable damages. Moreover, in some other scenarios, network lifetime is considered as the period until the entire sensing region is covered (Fanian & Kuchaki, 2019).

In homogeneous WSNs, as formerly appointed, nearby sensors usually have correlated data. Thus, freezing of some sensor nodes is not critical, and the network is reliable as long as at least the determined number of nodes are active (Shokouhifar et al., 2015). In this paper, different definitions of the network lifetime such as FND, Half Nodes Die (HND) and Last Node Dies (LND), throughput, minimum energy of the network versus rounds, are used to assess the performance of the proposed FSFA.

5. OUR PROPOSED MODEL

In the following sections, authors present procedures for the fitness function calculation and population vector initialization followed by hybrid-clustering algorithm proposition.

5.1 FITNESS FUNCTION CALCULATION

Optimal CH selection is deliberated as an optimization problem (Zenga & Dong, 2016). To manage the dissipated energy aiming at maximizing the network lifetime of a clustered WSN, most eligible CHs should be qualified based on a fitness value calculated by a fitness function for each node (Zenga & Dong, 2016). The energy parameters of the sensor nodes ensure that nodes with greater energy are given higher priority in the CH selection process. In this part, a multi-objective fitness function is formulated in order to evaluate feasible solutions for each clustering algorithm. This multi objective function is formulated as a weighted average of the five objective functions in Eq. (8) (Barzin et al., 2019). It comprises of five parameters, which will be discussing by the end of this section.

$$\mathrm{OF}_{SFFA} = \mathrm{Min}\left\{w_1 \times f_1 + w_2 \times f_2 + w_3 \times f_3 + w_4 \times f_4 + w_5 \times f_5\right\}$$

$$subject\ \text{to } 0 \le w_i \le 1, \sum_{i=1}^{5} w_i = 1,\ \mathrm{i} \in \{1,2,3,4,5\} \quad (8)$$

Suppose an AOI (area of interest) with the dimension of $M \times M$ in which N nodes randomly distributed. If there are K predefined clusters each having $|CHj|$ number of nodes, the average Euclidean distance of nodes $d_{ave}(j)$ to their associated cluster head CH_j is defined as :

$$f(1) = d_{ave}(j) = \frac{\frac{1}{N}\sum_{i=1}^{N} d(n_i, CH_j)}{A/2}, \ j = 1, 2, \cdots, k \quad (9)$$

In which, $d(ni,CHi)$ is the distance between node n_i and all cluster heads CH_j. The maximum of this average distance should be minimized. Additionally, the ratio of total initial energy of all nodes with the total current energy of the cluster head candidates in the current round E_r is defined as:

$$f(2) = E_r = \frac{\sum_{i=1}^{N} E(n_i)}{\sum_{j=1}^{k} E(CH_j)}, \quad j = 1, 2, \cdots, k \quad (10)$$

To increase the accuracy and to eliminate the non-clustered nodes from the process of clustering, the overlap ratio is added as the third fitness parameter, which is defined as follows:

$$f(3) = Overlap_{CH} = \frac{(N-k) - \sum_{j=1}^{k} \left|CH_j\right|}{\sum_{j=1}^{k} \left|CH_j\right|}, \ j = 1, 2, \cdots, k \quad (11)$$

Where N is the total number of sensor nodes, K is the number of CHs, and $|CHj|$ is the number of nodes that belong to cluster CH_j. The average Euclidean distance of cluster head to the sink is considered as the fourth parameter of the fitness function defined as:

$$f(4) = \frac{\frac{1}{k}\sum_{j}^{k}(CHj\,,\,\mathrm{d}_{\mathrm{Sin}k})}{M/2}\,,\; j = 1,2,..,k \tag{12}$$

In which d_{Sink} is the distance from each CHj to the sink. For large-scale networks, this distance should be kept minimized; otherwise, the energy some nodes will be lost. Whereas, for a small-scale network, that has a few closely located nodes, direct transfer of nodes to sink may be an acceptable option. Finally for the last parameter, *f(5)* is defined to maximize the average inter-cluster distances between CHs, can be expressed as Eq. (14), where CH_k is the nearest CH to CH_j, and the term A is used to normalize *f(5)*.

$$f(5) = \frac{A}{\frac{1}{|C_j|}\sum_{j=1}^{|C_j|} d\left(CH_j, CH_k\right)}\,, j = 1,2,\ldots,\ \mathrm{k} \tag{13}$$

Here, the normalization terms used for *f(1)*, *f(4)*, and *f(5)* are explained through an example. Suppose a symmetrically distributed network area (AOI) of dimension $M{\times}M$ with C number of clusters. If A be the average distance of two neighbor CHs, distances denoted by Eq. (9) and (13) can be normalized via $A/2$ and A, respectively. Moreover, the average distances of CHs to the sink in Eq. (10) can be normalized using $M/2$. The value of the parameter A can be approximately calculated as $A = M/\sqrt{C}$. For example, assuming M=100 and C=16, A can be calculated as A=100/4=25. Thus, each CH approximately covers an area of dimension 25×25 (Barzin et al., 2019).

5.2 ENCODING OF INDIVIDUALS

In the proposed approach, each individual of the population can be denoted as a binary string of length N, where N is the number of alive nodes. Value of “0” indicates a member node and value of “1” indicates a CH node. This structure is used for the representation of feasible solutions in both SFLA and FFA. An example of encoding a feasible solution can be shown in Figure 2. It should be mentioned that because of the continuous characteristics of the population updating process in both algorithms, the solutions have continuous values between 0 and 1. Therefore, only in the fitness evaluation phase, the solutions are rounded into binary structures in order to calculate the objective function. Afterward, the solutions are updated in the continuous structures.

1	*2*	*3*	*4*	*5*	*6*	...	*N*
1	1	0	0	1	0	…	1

Figure 2. Encoding of the individuals

5.3 HYBRID CLUSTERING ALGORITHM

In terms of evolutionary algorithms (EAs), hybridization is done mainly to improve performance by reaching better sets of solutions (Al-Ghazzali, 2002). SFLA is a collaborative population-based algorithm with high computational efficiency and good global search capability that can solve both discrete and continuous optimization problems (Xunli & Feieif, 2015). FFA, on the other hand, belongs to the group of stochastic algorithms and focuses on producing solutions at the lowest level within a search space. Its random search avoids falling into the premature local optimal. This algorithm has several advantages over other meta-heuristic algorithms. The firefly algorithm is based on absorption and brightness. This will automatically divide the entire population into subgroups with a mean interval, which helps each group crowd around local optima. Among all these local optima, the best global optimum could be attainable. Additionally, this classification allows for an improved search for all nonlinear multi-level optimization problems. The algorithm is set to a ratio of repetition so that convergence can be speeded up by leveraging these parameters. These benefits are faced with clustering, classifications and hybrid optimization (Fister et al., 2013).

The proposed algorithm will take advantage of both algorithms while trying to minimize any substantial disadvantage at the same time (Zhang et al., 2016). Moreover, the Rolette Wheel Selection (RWS) strategy is added to compensate for the probable lack of exploration (Yang X.Sh. 2010). At each iteration of the suggested SFSA, during the first step, an initial population of individuals P0 is generated random uniformly. In the second step, P0 is divided into subpopulations P1 and P2 for FFA and SFLA individuals. Through the third step, each individual's fitness value is calculated based on the proposed multi-objective fitness function presented in Eq. (9). In the fourth step, by the execution of a predefined number of FFA and SFLA based evolutionary processes, P1 and P2 are evolved paralleled. In the fifth step, the global bests of the FFA and SFLA algorithm's population are updated. Then in the sixth step, populations of P1 and P2, are shuffled and rearranged to new P1 and P2 and randomly selected for the next iteration using the RWS. Steps of SFSA are repeated until the maximum number of iterations is reached (Barzin et al., 2019).

5.3.1 TIME COMPLEXITY ANALYSIS

Time complexity analysis of FSFA is summarized in Table 2. The CH selection via FSFA at each round has a time complexity of $O(MaxIter \times PopSize \times N \times C)$, in which N is the total number of nodes, C is the desired number of clusters, and MaxIter is the number of algorithm's iterations. Moreover, *PopSize* is the population of the FSFA, which can be calculated as $PopSize = PopSFLA + PopFFA$, in which *PopSFLA* and *PopFFA* are the population size of SFLA and FFA, respectively. Furthermore, suppose N_{it} be the number of iterations for clustering main loop, N_{pop} be the population size of the swarm and TC-fitness is the time complexity of the cost function, presented by $TC_{fitness} = C \times N_{pop}$. It can be concluded that, the suggested FSFA, would not negatively impact on the time complexity of the clustering algorithm because it is linear in terms of both numbers of iteration and the population size. Thus, the hybrid algorithm is preferred, if it achieves better performance than the two constituents do.

Table 2. Time complexity for FFA, SFLA and FSFA

Algorithm	Time complexity
SFLA	$O(N_{it} \times N_{pop\_frog} \times \mathrm{Cost})$
FFA	$O(N_{it} \times N_{pop\_firefly} \times \mathrm{Cost})$
FSFA	$O(N_{it} \times (N_{pop\_frog} + N_{pop\_firfly}) \times TC_{fitness})$

6. SIMULATION AND RESULTS

In this section, assumptions of simulation containing the performance indices, algorithm and network parameters and, simulation result using MATLAB is presented, to evaluate the performance of the proposed algorithm concerning energy efficiency and network lifetime.

6.1 PERFORMANCE INDICES

Several indices for energy efficiency and network lifetime longevity have already been introduced in literature to evaluate the performance of the clustering protocols in WSNs (Gupta & Jha, 2018). Amongst the different indices, those used in this paper are listed in Table 3.

Table 3. Definition of the performance indices used in this paper.

Performance indices	Description
First Node Dead (FND)	Number of rounds in which the first sensor node of the network dies.
Half Node Dead (HND)	Number of rounds in which half number of sensor nodes are dead.
Last Dead Node (LND)	Number of rounds in which all sensor nodes of the network are dead.
Minimum Residual Energy Per Round (MREPR)	Total minimum residual energy of the network versus rounds.
Number of Alive Nodes Per Round (NANPR)	This reflects the total number of alive sensor nodes versus rounds.
Throughput	Measures the total No. of data packets successfully received at the sink versus rounds.

6.2 ALGORITHM AND NETWORK PARAMETERS

Adjustment of the controllable parameters of metaheuristic algorithms is of the highest importance before examining these techniques. To achieve this purpose, diverse values were explored and the most fitted ones were selected for each parameter. Parameter values for proposed FSFA including the parameters of the basic algorithms, and the multi-objective function's parameters of clustering can be summarized in Table 4. The weights of the multi-objective functions for clustering in Eq. (9) were tuned to maximize the FND, because in almost all applications FND is the most important measure (Shokouhifar, 2015). However, as previously mentioned, the proposed FSFA has an application-specific approach, which can be adapted with any application through the regulation of five weights expressed in Eq. (9) (Barzin et al., 2019). Simulations were carried out for WSNs in two scenarios, each comprise of five WSNs of size: 30, 60, 90,120 and 150 sensor nodes randomly deployed in topological areas of dimension 100 m ×100 m. In the first scenario the sink is situated at the center of the sensing area where as the in the second scenario the sink is positioned outside. All nodes have identical initial energy. The simple radio energy dissipation model (refer to Sec. 4.1) is used for all communications. Data aggregation is done according to the aggregation proposed by Oladimeji et al. (2016), where the aggregation rate, R_{agg} is set to 0.3. To eliminate the experimental error caused by randomness, each experiment run several times for each WSN. In order to adjust the controllable parameters of FFA and SFLA, different values were assessed for each parameter, and the most proper values were chosen. Appropriate parameter values for the proposed FFA, are mainly accustomed based on the guidelines offered by Mo, Ma, and Zheng (2013). Accordingly, a reasonable range for α is placed in [0.1, 0.2], additionally, too small or too big value of γ is not preferred, the optimal values for γ is located in [0.01, 30]. Proper value of β depends on the value of both β_0 and β_{min}, which are suggested to be 1 and 0.2 respectively. For simpler problems, a population size (*PopSize*) of 20 to 40 fireflies may be appropriate. However, when the problem becomes complicated, it should not be larger than 50. Parameter values for the proposed SFLA are set in accordance with the Nelder-Mead Standard and the ranges examined and advised by Wang and Gong (2013). More assumptions related to WSNs' are listed in Table 5.

Table 4. Parameters of basic algorithms

Parameter	Description	Value
MaxIter	Maximum Number of Iterations	50
PopSize (Pop_{SFLA}+Pop_{FFA})	Population Size	20 (10+10)
α in Eq. (4)	Random Change Coefficient in FFA (Eq. 3)	0.1
β in Eq. (4)	Light Absorption Coefficient in FFA (Eq. 3)	0.5
γ in Eq. (4)	Attraction Coefficient (Eq. 3)	0.05
nMemeplex	Number of frogs in each memeplex for SFLA	5
nPopMemeplex	Nelder-Mead Standard number of memplexes for SFLA	2
w_1 , w_2 , w_3 , w_4 ,w_5	Weights of Multi-Objective Function in Clustering (Eq.9)	0.5 , 0.2 , 0.1 , 0.1, 0.1

Table 5. WSNs' network parameter assumptions

Network Parameters	Value
Area of Interest ($M \times M$)	(100 m × 100 m)
Number of sensor nodes (N) scenario #1 (Sink at the Center)	30, 60, 90, 120, 150
Number of sensor nodes (N) scenario #2 (Sink at the Top Border)	30, 60, 90, 120, 150
Maximum No. of Rounds	1500
Number of clusters (C)	$0.1 \times N$
Initial energy of sensor nodes	0.3 J
E_{elec}	50 nJ/bit
E_{fs}	100 pJ/bit/m^2
E_{mp}	0.013 pJ/bit/m^4
d_0	87.0 m
E_{DA}	5 n J/bit
Packet Size	4000 bits
R_{agg}	0.3

6.3 SIMULATION RESULTS

6.3.1 COMPARISION WITH THE BASIC ALGORITHMS

Since FSFA is a hybrid algorithm, it should be first compared with its constituents SFLA and FFA algorithms, in terms of FND, HND and LND. To achieve that, three metaheuristic algorithms utilize the same clustering process with the same parameters, which can be reviewed in Table 4. In order to have a fair comparison between the different techniques with the same time complexity, population size of SFLA and FFA were set at 20. Simulation results in two scenarios mentioned in Table 4, can be summarized in Table 6 and 7 in terms of FND, HND, LND and throughput respectively. Additionally, Figure 6 and 7 statistically qualifies them for the NANPR and MREPR.

Table 6. Comparison of FSFA with SFLA and FFA for all WSNs, in terms of FND, HND, and LND

Scenario #1	WSN #1			WSN #2			WSN #3			WSN #4			WSN #5		
	FND	HND	LND	FND	HND	LND	FND	HND	LND	FND	HND	LND	FND	HND	LND
Proposed SFLA	639	644	648	651	669	681	587	677	787	515	681	962	483	684	1053
Proposed FFA	636	643	649	647	670	683	633	681	696	566	656	748	551	652	825
Proposed FSFA	646	657	662	677	690	701	686	698	711	643	701	722	579	710	779
Scenario #2	WSN #1			WSN #2			WSN #3			WSN #4			WSN #5		
	FND	HND	LND	FND	HND	LND	FND	HND	LND	FND	HND	LND	FND	HND	LND
Proposed SFLA	551	568	592	465	580	743	416	563	868	272	596	966	338	594	1082
Proposed FFA	545	533	612	512	537	793	464	530	792	397	558	812	353	568	756
Proposed FSFA	564	579	618	516	584	854	483	606	733	401	627	745	455	624	775

Table 7. Comparison of FSFA with SFLA and FFA for all WSNs, in terms of Throughput

Scenario #1	WSN #1			WSN #2			WSN #3			WSN #4			WSN #5		
	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND
Proposed SFLA	19170	19244	19276	39060	39906	40061	52830	59829	60849	61800	77724	81806	72450	96938	103119
Proposed FFA	19080	19221	19255	38820	39932	40017	56970	60536	60818	67920	77778	81339	82650	96670	102581
Proposed FSFA	19380	19638	19659	40620	41235	41311	61740	62535	62727	77160	83398	83707	86850	104118	105440
Scenario #2	WSN #1			WSN #2			WSN #3			WSN #4			WSN #5		
	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND
Proposed SFLA	16530	16925	17085	27900	33255	34447	37440	48184	53094	32640	66263	72036	50700	82461	91070
Proposed FFA	16350	15960	17104	30840	32079	34641	41760	47079	53337	47640	65563	72550	52950	82427	91086
Proposed FSFA	16920	17230	17437	30960	34027	35322	43470	51305	54341	48120	70851	73940	68250	87900	92780

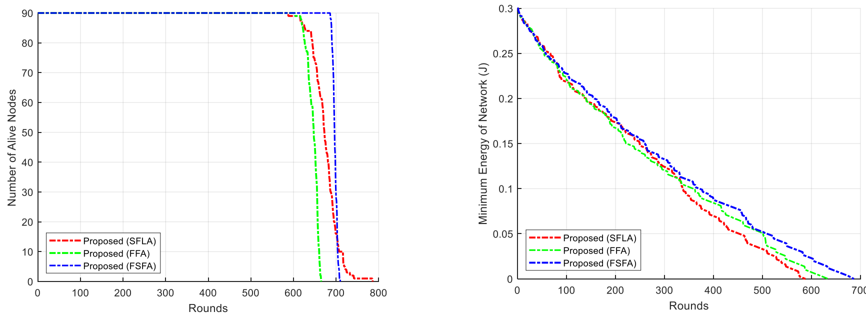

Figure 3. Comparison of the NANPR and MREPR (for Scenario #1) for the proposed SFLA, FFA and FSFA

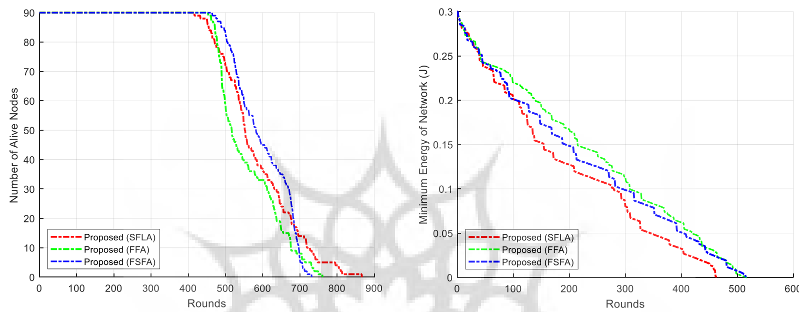

Figure 4. Comparison of the NANPR and MREPR (for Scenario #2) for the proposed SFLA, FFA and FSFA

6.3.2 COMPARISION WITH THE EXISTING PROTOCOLS

In this section, two scenarios with different number of sensor nodes in the same network sizes are simulated to evaluate the performance of FSFA comparing with the existing protocols. The network size remains fixed on emphasizing a single hopping scheme (Each node sends its data to the sink via one CH). All other simulation parameters including weights of multi-objective Function in clustering are derived from Table 4 and 5. There are four existing protocols used in the simulation, including LEACH, ERA, SIF and FSFLA. These protocols were selected to make the comparison more reasonable via diversifying the performance evaluation of the proposed FSFA against a miscellaneous range of clustering approaches including classical, fuzzy, metaheuristic and hybrid based methods. Furthermore, they had studied during the literature review and all supports the single hopping scheme in their initial version. To see the results, Table 8 contains values of FND, HND and LND for all five protocols in different sizes. Table 9 shows the throughput values of each network protocols. Additionally, Figure 8 and 9 statistically evaluate them in terms of NANPR and MREPR for each scenario respectively. Similar to section 6-3-1, the performance of the hybrid model against the existing protocols depends on the values calculated for FND.

Table 8. Comparison of the lifetime indices (FND, HND, and LND) between FSFA and existing protocols in both scenarios

Scenario #1	WSN #1			WSN #2			WSN #3			WSN #4			WSN #5		
	FND	HND	LND	FND	HND	LND	FND	HND	LND	FND	HND	LND	FND	HND	LND
LEACH	537	632	750	485	670	799	546	680	812	536	684	855	505	703	819
ERA	409	521	646	491	577	669	524	590	669	495	585	697	495	606	694
SIF	566	576	597	589	608	657	613	627	673	607	628	686	596	641	672
FSFLA	588	594	599	589	605	613	591	608	615	580	604	615	582	618	628
Proposed FSFA	646	657	662	677	690	701	686	698	711	643	701	722	612	714	785
Scenario #2	WSN #1			WSN #2			WSN #3			WSN #4			WSN #5		
	FND	HND	LND	FND	HND	LND	FND	HND	LND	FND	HND	LND	FND	HND	LND
LEACH	360	558	701	311	536	806	340	579	802	377	594	870	356	604	902
ERA	220	467	612	369	512	608	392	521	640	380	530	667	431	554	649
SIF	465	531	564	449	535	572	453	567	796	469	593	630	411	594	652
FSFLA	485	543	570	506	536	547	464	564	586	401	565	591	353	574	586
Proposed FSFA	564	579	613	514	584	854	493	606	733	508	627	745	490	624	756

Table 9. Comparison of throughput values of each network protocols for both scenarios

	WSN #1			WSN #2			WSN #3			WSN #4			WSN #5		
Scenario #1	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND
LEACH	16110	18480	19073	29100	38513	39979	49140	58999	61215	64320	78872	82096	75750	100728	104292
ERA	12270	14851	15756	29460	33483	34272	47160	51929	53444	59400	67913	70296	74250	88266	90554
SIF	16980	17199	17298	35340	36220	36611	55170	56117	56558	72840	74839	75541	91800	95306	96231
FSFLA	17640	17773	17818	35340	36137	36286	53190	54473	54675	69600	72033	72443	87300	91982	92457
Proposed FSFA	19380	19638	19659	40620	41235	41311	61740	62535	62727	77160	83398	83707	86850	104118	105440
	WSN #1			WSN #2			WSN #3			WSN #4			WSN #5		
Scenario #2	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND	THR_FND	THR_HND	THR_LND
LEACH	10800	15239	16023	18660	29684	32731	30600	47465	51278	45240	66063	71164	53400	83474	90047
ERA	6600	12711	13695	22140	29250	30544	35280	44816	47211	45600	61191	64916	64650	80119	83020
SIF	13950	15120	15397	26940	31059	31683	40770	47761	50035	56280	67283	68529	61650	84118	86254
FSFLA	14550	16043	16298	30360	31828	32035	41760	48980	49651	48120	66744	67602	52950	84902	85455
Proposed FSFA	16920	17230	17437	30840	34027	35322	44370	51305	54341	60960	70851	73940	73500	87900	92780

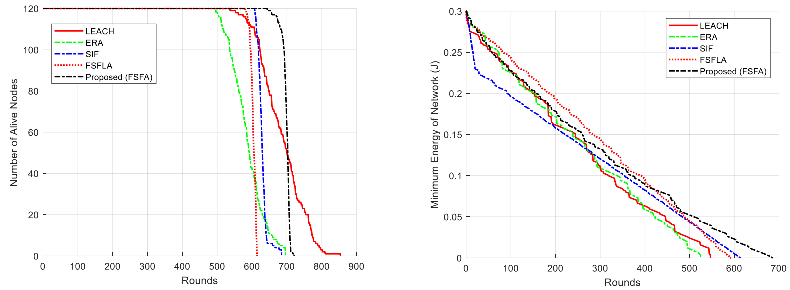

Figure 5. Comparison of the NANPR and MREPR for five WSNs (in average) of scenario #1

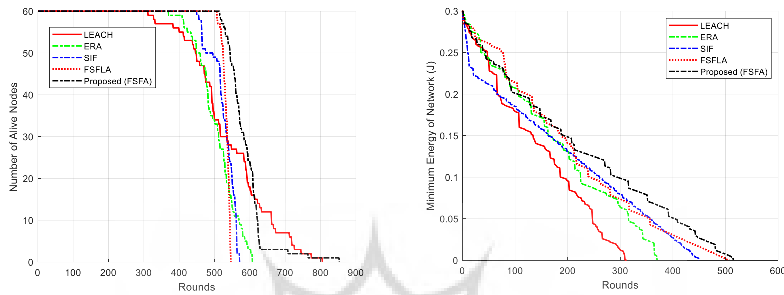

Figure 6. Comparison of the NANPR and MREPR for five WSNs (in average) of scenario #2

6.3.4 DISCUSSION

In this section, in order to justify the outperformance of the proposed FSFA, the average values acquired for all WSNs in two scenarios will be discussed in terms of energy efficiency and network lifetime. Considering the mean values in Figure 9, the proposed FSFA algorithm can prolong FND by 25.11%, 35.21%, 9.86%, and 11.4% compared with LEACH, ERA, SIF, and FSFLA respectively for the first scenario (Sink placed at the center of the area of interest). Moreover, FSFA similarly improves the mean values of FND by 47.31%, 43.36%, 14.32%, and 16.30% compared with LEACH, ERA, SIF, and FSFLA correspondingly for the second scenario (Sink placed at the top border of the area of interest). Likewise, it slightly improves the mean values of LND comparing with ERA, SIF, and FSLFA for both scenarios. Similarly, FSFA improves the mean value of FND throughput depicted in Figure 10 denoted by TRU_FND by 21.91%, 28.40%, 5.0%, and 8.62% compared with LEACH, ERA, SIF, and FSFLA respectively for the first scenario. Likewise, FSFA improves the mean value of FND throughput by 42.78%, 30.02%, 13.53%, and 20.70% compared with LEACH, ERA, SIF, and FSFLA respectively for the second scenario. Finally, the total improvement of FND mean values, regardless of the position of the sink, are 33.95%, 32.62%, 12.09%, and 13.85 compared with LEACH, ERA, SIF, and FSFLA respectively. As for the total FND throughput mean values improvements are 33.3%, 29.21%, 9.27%, and 14.66% compared with LEACH, ERA, SIF, and FSFLA respectively. Amongst all examined clustering approaches SIF showed a better performance and got the second rank. FSFLA got the third rank. ERA got the fourth place and LEACH had the lowest performance.

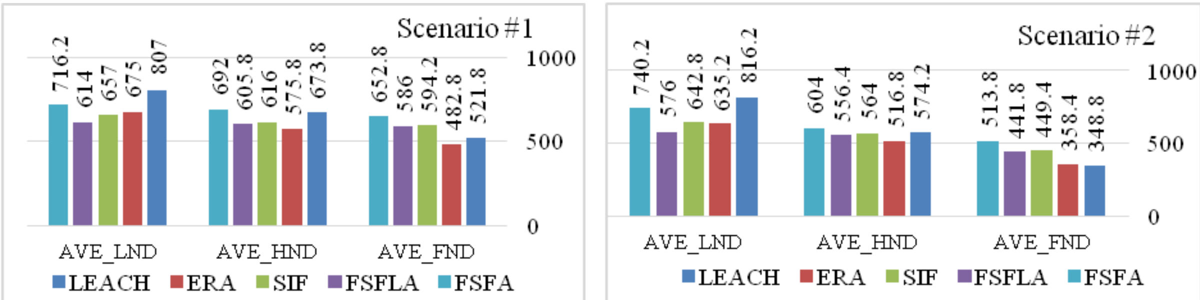

Figure 7. Comparison of the FND (on average) for all WSNs in two scenarios

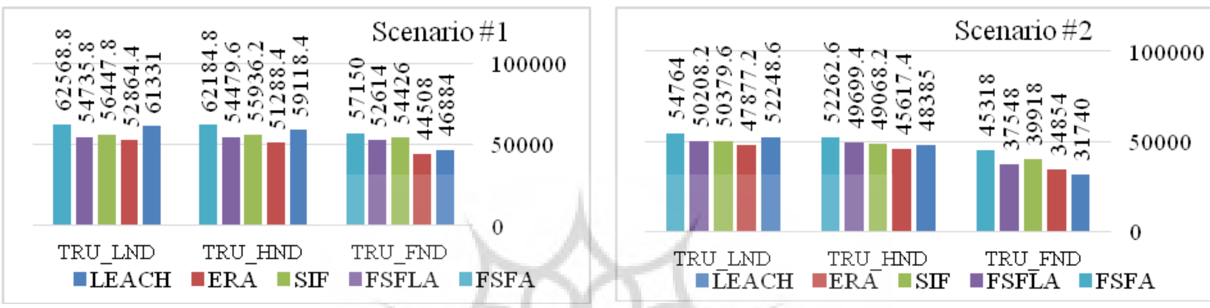

Figure 8. Comparison of the Throughputs (on average) for all WSNs in two scenarios

As it can be inferred from the results, regardless of the position of the sink, (either inside or outside the sensing field) the proposed hybrid model is capable of enlarging the life span of a broad range of WSNs from small to medium scale for single-hop clustering-based routing scheme compared with the existing protocols in an application-specific way. This is done via the effective CH selection technique (FSFA) and managing the dissipated energy of the nodes more efficiently leading to maximizing the FND. FND is considered the most important metrics to evaluate the performance WSNs in the majority of applications. Although, the proposed model has the ability to cope with other application through manipulating weights of Eq. (9) which improves other indices such as HND or LND, according to application requirements.

7. CONCLUSION

In this paper, a new hybrid clustering Algorithm (FSFA) was proposed which blends two basic bio-inspired based metaheuristic algorithms including the firefly algorithm (FFA) and shuffled frog leaping algorithm (SFLA) aiming at energy efficiency and prolonging network lifetime of the WSNs. As presented in section 6 energy consumption of WSNs as well as the network lifetime simulated with MATLAB 2018a in two different scenarios. Results indicate a meaningful performance improvement of the proposed FSFA not only over its constituents SFLA and FFA in terms of energy dissipation and network lifetime but also against existing protocols including LEACH, ERA, SIF and FSFLA especially in terms of FND. FSFA protocol was mainly focused on the clustering algorithm. Naturally, the routing phase of the protocol was concentrated for the next and further studying. The authors propose extending the study by focusing on the steady phase, optimizing the data aggregation and advertisement approaches in the future.

References

Abba Ari, A., Yenke, B.O., Labraoui, N., Damakoa, I., & Gueroui, A. (2016). A power efficient cluster-based routing algorithm for wireless sensor networks: Honeybees swarm intelligence-based approach. *Journal of Network and Computer Applications, 69*, 77-97.

Abbasi, A.A., & Younis, M. (2007). *A survey on clustering algorithms for wireless sensor networks", Computer Communications*, Vol. 30, 2826–2841.

Al-Ghazzali, T (2002). A taxonomy of hybrid metaheuristics. *Journal of Heuristics, 8*, 541–564.

Al-Ghazzali, T. (2009). *Metaheuristics: from design to implementation*. Chichester: John Wiley and Sons Inc.

Amgoth, T., & Jana, P. K. (2015). Energy-aware routing algorithms for wireless sensor networks. *Computers & Electrical Engineering, 41*, 357-367.

Anandamurugan, S., & Abirami T. (2017). Antipredator adaptation shuffled frog leap algorithm to improve network lifetime in wireless sensor network. *Wireless Personal Communications, 94*, 2031–2042.

Bagci, H., & Yazici, A. (2013). An energy aware fuzzy approach to unequal clustering in wireless sensor networks, *Applied Soft Computing*, 13(4), 1741-1749.

Barzin, A., Sadeghieh, A., Khademi Zareh, H., & Honarvar, M. (2019). Hybrid swarm intelligence-based clustering algorithm for energy management in wireless sensor networks. *Journal of Industrial and Systems Engineering, 12(3), 78-106.*

Butenko, V., Nazarenko, A., Sarian, V., Sushchenko, N., & Lutokhin, A. (2014). Applications of wireless sensor networks in next generation networks, international telecommunication union, *Telecommunication Standardization Sector of ITU-T,* Series Y. 2000: Ngn-Awsn (2014-02).

Dietrich, I., & Dressler, F. (2009). On the lifetime of wireless sensor networks. *ACM Transactions on Sensor Networks, 5*, 1-38.

Eusuff, H., Lansey, M., & Pasha F. (2006). Shuffled frog-leaping algorithm: A memetic meta-heuristic for discrete optimization. *Engineering Optimization, 38*(2), 129-154.

Fanian, F., & Kuchaki Rafsanjani, M. (2019). Cluster-based routing protocols in wireless sensor networks: A survey based on methodology. *Journal of Network and Computer Applications, 142, 111-142*

Fanian, F., & Rafsanjani, M.K. (2018). Memetic fuzzy clustering protocol for wireless sensor networks: Shuffled frog leaping algorithm. *Applied Soft Computing, 71*, 568-590.

Fister, I., Fister, I., Yang, X., & Brest, J. (2013). *A Comprehensive Review of ifrelfy Algorithms, Swar□□ and Evolutionary Computation*, *13*, 34-46.

Gupta, G., & Jha, S. (2018). Integrated Clustering and routing protocol for wireless sensor networks using cuckoo and harmony search based metaheuristic techniques. *Engineering Applications of Artificial Intelligence, 68*, 101-109.

Hamzeloei, F., & KhalilyDermany, M. (2016). A topsis based cluster head selection for wireless sensor network. *Procedia Computer Science*, *98*, 8-15.

Hanifi, A., Taghva, M., Haghi, R.H., & Feizi, K. (2018). Clustering for reduction of energy consumption in wireless sensor networks by AHP method. *Journal of Information Systems and Telecommunication, 6*(1), 9-17.

Heinzelman, W., Chandrakasan, A., & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless micro sensor networks. In *Proceeding of the Hawaii International Conference on Systems Science*, Vol. 8.

Heinzelman, W., Chandrakasan, A. & Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communication*, 1(4).

Jabeur, N. (2016). A firefly-inspired micro and macro clustering approach for wireless sensor networks. *The Seventh International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN).*

Kassan, S., Gaber, J., & Lorenz P. (2018). Game theory based distributed clustering approach to maximize wireless sensors network lifetime. *Journal of Network and Computer Applications, 123*, 80-88.

Khorsandi, A., Alimardani A., Vahidi B., &. Hosseinian S.H. (2011). Hybrid shuffled frog leaping algorithm and nelder–mead simplex search for optimal reactive power dispatch. *IET Generation Transmission & Distribution, 5(2).*

Ko, A., Lau, Y.K., & Sham P.S. (2008). Application of distributed wireless sensor network on humanitarian search and rescue systems. *In Proceeding of the Second International Conference on Future Generation Communication and Networking*, 02, 328-333.

Minaie, A., & Sanati-Mehrizy, A. (2013). Application of wireless sensor networks in health care system. *In proceeding of the 120th ASEE annual conference & exposition*.

Mo, Y., Ma, Y., & Zheng, Q. (2013). Optimal choice of parameters for firefly algorithm. IEEE, *Fourth International Conference on Digital Manufacturing & Automation, Qingdao*, 887-89.

Mukhdeep, S. M., & Singh, S.B. (2016). Firefly algorithm based clustering technique for wireless sensor networks. *Wispnet Conference, IEEE Press*.

Oladimeji, M.O., Turkey, M., & Dudley, S. (2017). HACH: Heuristic algorithm for clustering hierarchy protocol in wireless sensor networks. *Applied Soft Computing, 55*, 452-461.

Pantoni, R.P., & Brandão, D. (2013). A gradient based routing scheme for street lighting wireless sensor networks. *Journal of Network and Computer Applications*, 36(1), 77-90.

Pratyay, K., & Prasanta, K.J. (2014). Energy efficient clustering and routing algorithms for wireless sensor networks: Particle swarm optimization approach. *Engineering Applications of Artificial Intelligence*, *33*, 127-140.

Ran, G., Zhang, H., & Gong, S. H. (2010). Improving on LEACH protocol of wireless sensor networks using fuzzy logic. *Journal of Information and Computational Science*, *7*, 767-775.

Shokouhifar, M., & Jalali, A. (2015). A new evolutionary based application specific routing protocol for clustered wireless sensor networks. *AEU - Electronics and Communications*, *69*, 432-441.

Singh, M.P., & Singh, S. (2017). Improved metaheuristic based energy-efficient clustering protocol for wireless sensor networks. *Engineering Applications of Artificial Intelligence*, *53*, 142-152.

Sohraby, K., Minoli, D., & Znati, T. (2007). *Wireless sensor networks technology, protocols, and applications*. John Wiley & Sons Ltd.

Tripathi, M., Gaur, M.S., Laxmi, V., & Battula, R.B. (2013). Energy efficient LEACH-C protocol for wireless sensor networks. *Third International Conference on Computational Intelligence and Information Technology (CIIT 2013).*

Wang, L., & Gong, Y. (2013). Convergence and parameters analysis of shuffled frog leaping algorithm. *International Conference on Artificial Intelligence and Software Engineering (ICAISE).*

Xunli, F.A, & Feieif, D.U. (2015). Shuflfed frog leaping algorithm based unequal clustering strategy for wireless sensor networks. *International Journal of Applied Mathematics and Information Sciences*, *9*, 1415–1426.

Yang, X. Sh. (2010). *Nature-inspired metaheuristic algorithms*. Luniver Press.

Zahedi, Z.M., Akbari, R., Shokouhifar, M., Safaei, F., & Jalali, A. (2016). Swarm intelligence based fuzzy routing protocol for clustered wireless sensor networks. *Expert Systems With Applications*, *55*, 313-328.

Zenga, B., & Dong, Y. (2016). An improved harmony search based energy-efficient routing algorithm for wireless sensor networks. *Applied Soft Computing*, *41*, 135-147.

Zhang, L., Liu, L., Yang, X. Sh., & Dai, Y. (2016). A novel hybrid firefly algorithm for global optimization. *PLOS ONE*, 11(9), e0163230.

Bibliographic information of this paper for citing:

Barzin, Amirhossein; Sadegheih, Ahmad; Khademi Zare, Hassan, & Honarvar, Mahboobeh (2019). Hybrid bio-inspired clustering algorithm for energy efficient wireless sensor networks. *Journal of Information Technology Management*, 11(1), 76-101.